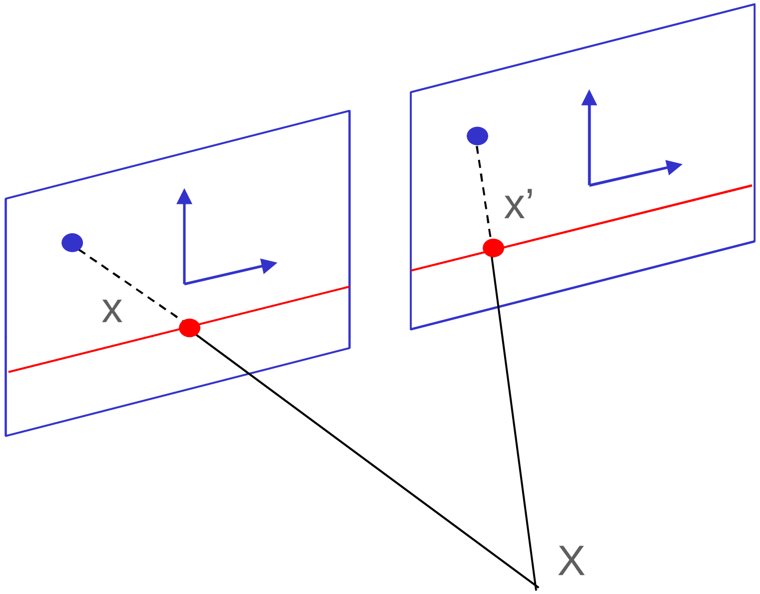
Stereo Reconstruction

problem

Goal: reconstruct the 3D shape of objects in the scene from 2 or more images.

special case



model

$$x_1 = f\,\frac{X_1}{X_3} \qquad x_2 = f\,\frac{X_2}{X_3}$$

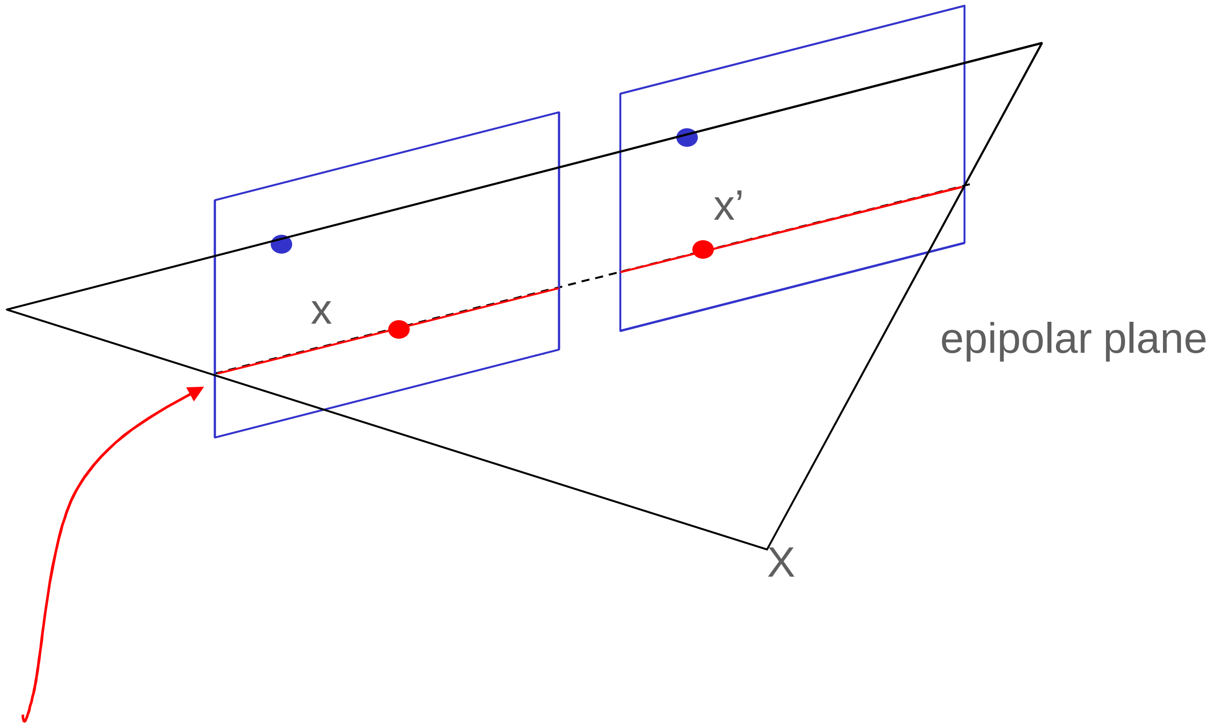
$$x'_1 = f\,\frac{X'_1}{X'_3} \qquad x'_2 = f\,\frac{X'_2}{X'_3}$$

$$X'_1 = X_1 - b \quad X'_2 = X_2 \quad X'_3 = X_3$$

reconstruction

$$X_3 x'_1 = X_3 x_1 - fb, \qquad X_3 = f\,\frac{b}{x_1 - x'_1} \quad \Rightarrow \quad X_3 = f\,\frac{b}{d}\,, \qquad d = x_1 - x'_1$$

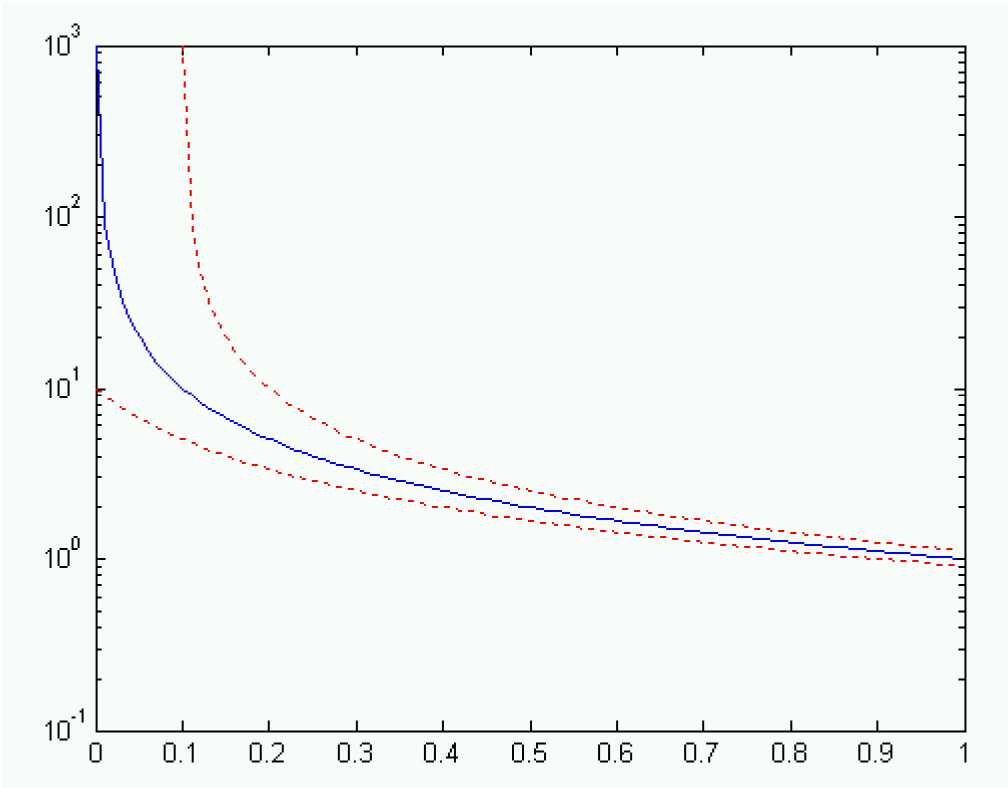
epipolar geometry



$$x_2 = x'_2$$

depth

X_3



normalized disparity (df/b)

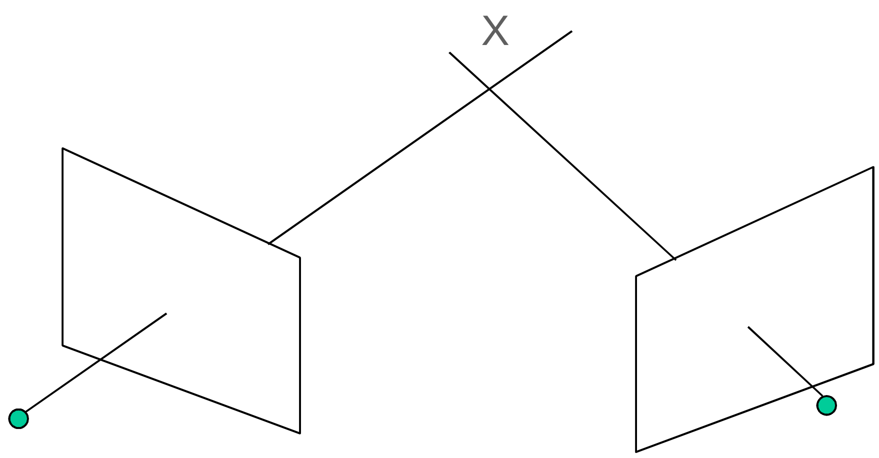
depth estimates obtained from the true disparity value and from noisy disparity values with 0.001 errors.

steps

Stereo reconstruction involves 2 steps.

matching: association of corresponding points in two (or more) images.

triangulation:



calibrated cameras

reconstruction

problem: reconstruct *X*, from the projections *x, x'* and matrices *P, P'.*

known cameras: P, P'

$$camera\,1: \quad x_1 = \frac{p_1 \cdot \tilde{X}}{p_3 \cdot \tilde{X}} \qquad x_2 = \frac{p_2 \cdot \tilde{X}}{p_3 \cdot \tilde{X}}$$

$$camera\,2: \quad x'_1 = \frac{p'_1 \cdot \tilde{X}}{p'_3 \cdot \tilde{X}} \qquad x'_2 = \frac{p'_2 \cdot \tilde{X}}{p'_3 \cdot \tilde{X}}$$

$\tilde{X}$ - homogeneous coord.

$$\begin{cases} (p_3 \cdot \tilde{X}) x_1 = p_1 \cdot \tilde{X} & (p_3 \cdot \tilde{X}) x_2 = p_2 \cdot \tilde{X} \\ (p'_3 \cdot \tilde{X}) x'_1 = p'_1 \cdot \tilde{X} & (p'_3 \cdot \tilde{X}) x'_2 = p'_2 \cdot \tilde{X} \end{cases}$$

algebraic method

$$(p_3 \cdot \tilde{X}) x_1 = p_1 \cdot \tilde{X}$$

$$(p_3 \cdot \tilde{X}) x_2 = p_2 \cdot \tilde{X}$$

$$(p'_3 \cdot \tilde{X}) x'_1 = p'_1 \cdot \tilde{X}$$

$$(p'_3 \cdot \tilde{X}) x'_2 = p'_2 \cdot \tilde{X}$$

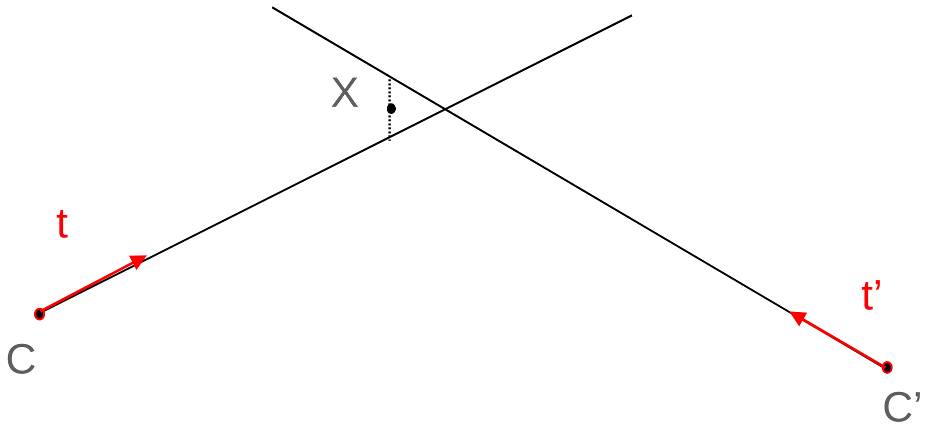
4 equations 3 unknowns

$$\boxed{M\tilde{X} = 0}$$

$$M = \begin{bmatrix} p_1^T - p_3^T x_1 \\ p_2^T - p_3^T x_2 \\ p'^T_1 - p'^T_3 x'_1 \\ p'^T_2 - p'^T_3 x'_2 \end{bmatrix}$$

Least squares: $\tilde{X}$ is a normalized eigenvector associated to the smallest eigenvalue.

geometric method



find the closest pair of points from both lines:

optical rays

$$X = C + \alpha t \qquad t = C - P^{\#} x$$

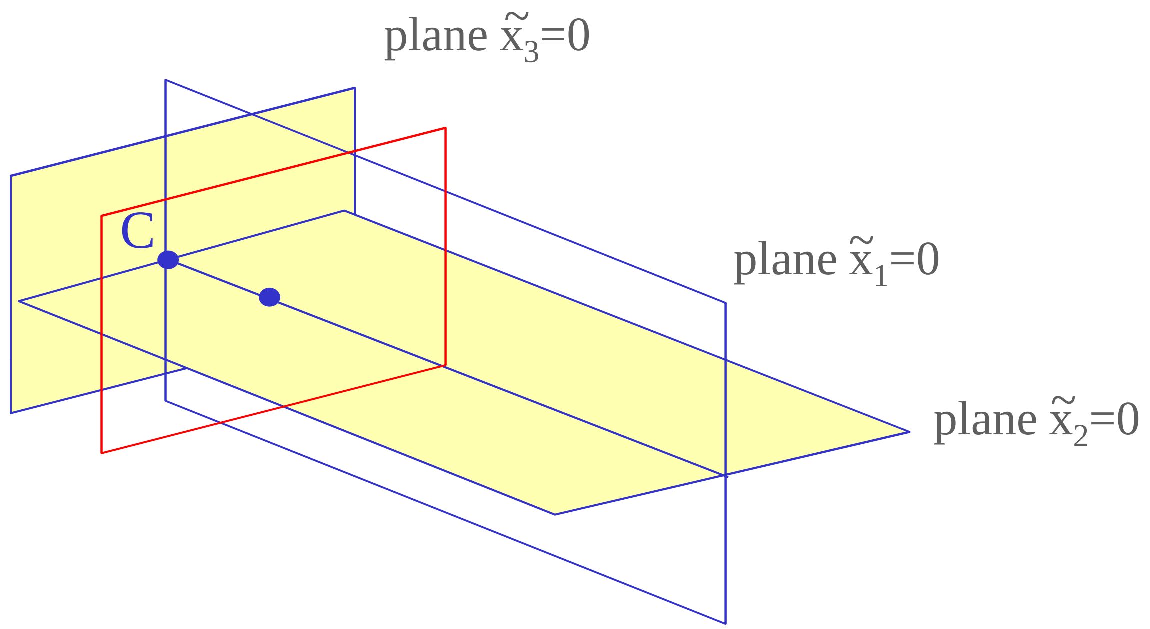
$$X' = C' + \alpha' t' \qquad t' = C' - P'^{\#} x'$$

$$\alpha = \frac{[t-(t^T t')t']^T (C-C')}{1-(t^T t')^2}$$

$$\alpha' = \frac{[t'-(t^T t')t]^T (C'-C)}{1-(t^T t')^2}$$

$$P^{\#} = P^T (PP^T)^{-1} \qquad P'^{\#} = P'^T (P' P'^T)^{-1}$$

optical center

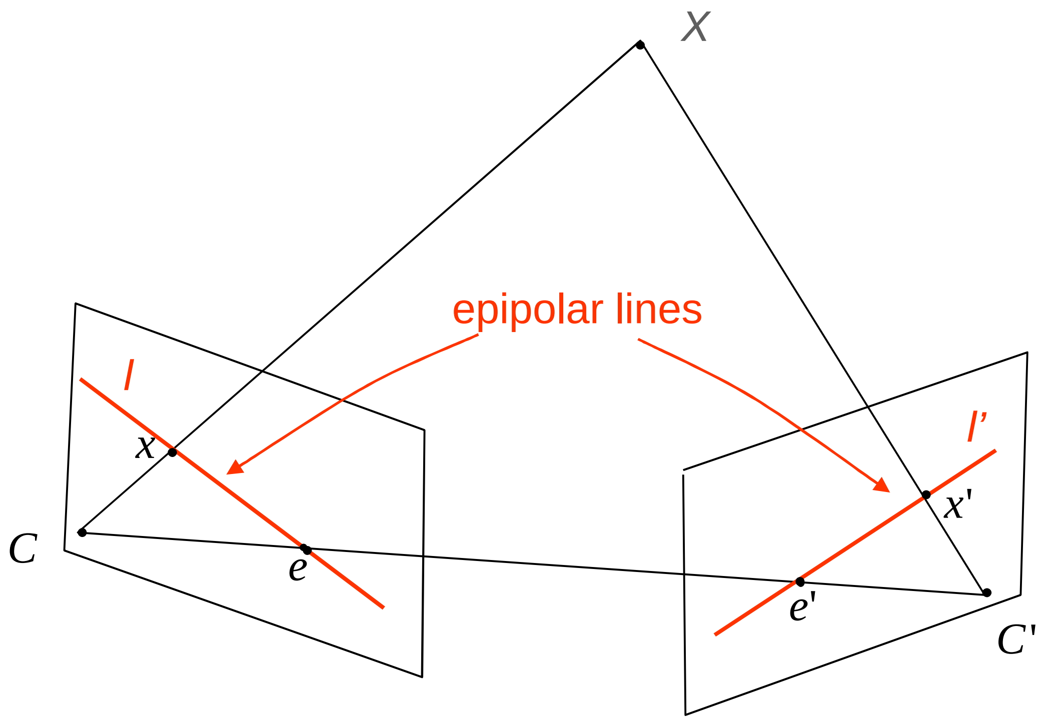


Centro óptico

$$P\tilde{C} = 0$$

epipolar geometry

epipolar geometry



XCC' –epipolar plane

l,l' – epipolar lines

e,e' – epipoles

epipolar geometry (2)

In general, two optical rays are non-coplanar and do not intersect. To guarantee that they meet the following condiction must hold

$$\tilde{x}'^T F \tilde{x} = 0$$

$\tilde{x}, \tilde{x}'$ – pair of corresponding points (homogeneous coord.)
F –fundamental matrix

The geometry of a pair of cameras is called epipolar geometry and it is characterized by F.

The epipolar geometry is used in both steps of stereo reconstruction.

Lines in 2D

A line

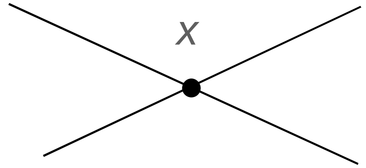
$$l_1 x_1 + l_2 x_2 + l_3 = 0$$

can be expressed as follows using homogeneous coordinates

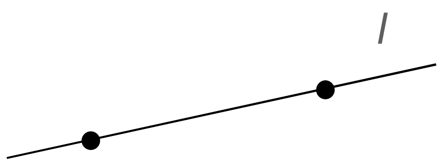
$$l^T \tilde{x} = 0 \qquad\qquad l \in R^3$$

There is a complete symmetry between points and lines in 2D using homogeneous coordinates.

Intersection of two lines: $\tilde{x} = l_1 \times l_2$



line defined by two points: $l = \tilde{x}_1 \times \tilde{x}_2$



Matrix representation of cross product

Cross product of two vectors

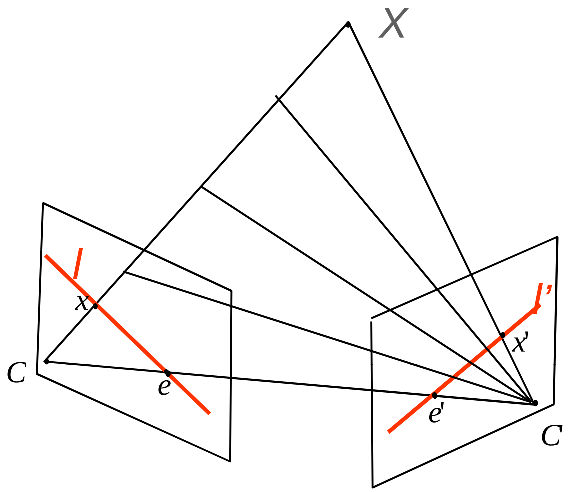
$$a \times b = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix}$$

Can be obtained by

$$a \times b = [a]_x b \qquad [a]_x = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

$[a]_x$ Is a skew symmetric matrix

Proof



hypothesis: calibrated cameras *(K=K'=I)* and world coordinate system centered in camara 1.

$$P=[I\ 0] \qquad P'=[R\ \ t]$$

1st step

Define 2 points belonging to the optical ray defined by C and *x*

$$\tilde{X}_A=\begin{bmatrix}0\\1\end{bmatrix} \qquad \tilde{X}_B=\begin{bmatrix}\tilde{x}\\0\end{bmatrix}$$

2nd step

Project them by the 2nd camera

$$\tilde{x}'_A=[R\ t]\begin{bmatrix}0\\1\end{bmatrix}=t \qquad \tilde{x}_B=[R\ t]\begin{bmatrix}\tilde{x}\\0\end{bmatrix}=R\tilde{x}$$

epipolar line

$$l=t\times R\tilde{x}=[\mathrm{t}]_\times R\tilde{x}$$

Proof (2)

3rd step

$\tilde{x}'$ belongs to the epipolar line

$$\tilde{x}'^T l = 0 \;\Rightarrow\; \tilde{x}'^T \underbrace{[t]_x R}_{E} \tilde{x} = 0$$

essential matrix

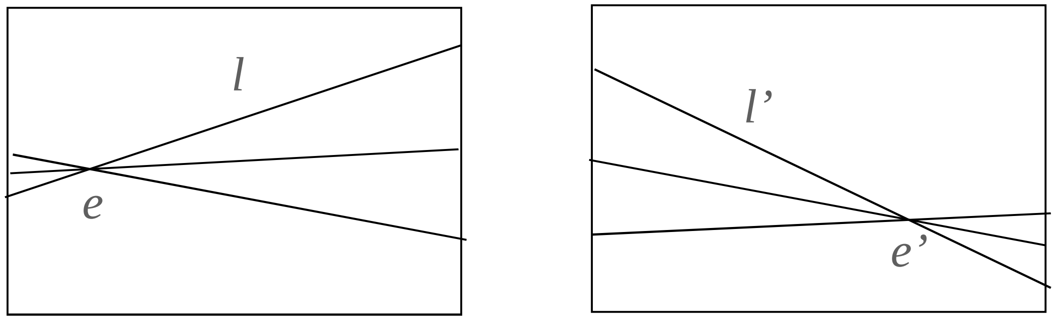
When the cameras are not calibrated we can replace

$$\tilde{x} \rightarrow K^{-1}\tilde{x}, \quad \tilde{x}' \rightarrow K'^{-1}\tilde{x}'$$

and obtain

$$\tilde{x}'^T F \tilde{x} = 0, \qquad F = K'^{-T}[t]_X R K^{-1}$$

epipoles e epipolar lines



epipolar lines $\quad l = F^T \tilde{x}', \qquad l' = F\tilde{x}$

epipoles $\quad F\tilde{e} = 0, \qquad F^{\mathrm{T}}\tilde{e}' = 0$

Propriedades da matriz fundamental

- *F* is a *3x3* matrix with rank *2 (det F =0)*
- fundamental property: $\tilde{x}'^T F\tilde{x} = 0$
- epipoles: $F\tilde{e} = 0, \qquad F^{\mathrm{T}}\tilde{e}' = 0$
- Epipolar lines: $l = F^T\tilde{x}', \qquad l' = F\tilde{x}$
- linhas epipolares $l' = F[k]_x l$ k – line which does not contain the epipole
- relationship with P, P' $F = [e']_x (P' P^{\#})$

Skew symmetric matrix $[a]_x = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$

uncalibrated cameras

Reconstruction

Matrix *F* can be obtained from pairs of corresponding points in both images.

F can be estimated even if the projection matrices *P, P*' are unknown

Can we reconstruct the scene knowing matrix *F* ?

Estimation of F

F can be estimated from 8 or more pairs of points (7 are enough).

Each pair of points defines a restriction

$$\begin{bmatrix} x'_{i1} & x'_{i2} & 1 \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} x_{i1} \\ xi_{2} \\ 1 \end{bmatrix} = 0$$

Restrictions can be expressed in the form

$$Mf = 0$$

$$M = \begin{bmatrix} x'_{11}x_{11} & x'_{11}x_{12} & x'_{11} & x'_{12}x_{11} & x'_{12}x_{12} & x'_{12} & x_{11} & x_{12} & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x'_{n1}x_{n1} & x'_{n1}x_{n2} & x'_{n1} & x'_{n2}x_{n1} & x'_{n2}x_{n2} & x'_{n2} & x'_{n2} & x_{n2} & 1 \end{bmatrix}$$

constraints: $||F||=1$ e $\det F = 0$

8 point algorithm

1. normalization: move the mean to the origin and normalize variance

$$\tilde{x}_i = T x_i \qquad \tilde{x}'_i = T' x'_i$$

2. Determine fundamental matrix

• least squares: $\tilde{f}$ is the eigenvector associated to the smallest eigenvalue of $M^T M$; $\tilde{F}$ is obtained from $\tilde{f}$

• rank 2: $\hat{F} = U \tilde{D} V^T$ where U, V are obtained by singular value decomposition of $\tilde{F}$ and

$$\tilde{D} = \begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

3. Denormalization

$$F = T' \hat{F} T$$

Theorem: the scene can be reconstructed from two uncalibrated cameras except for a projective transformation.

The reconstructed points $\hat{X}_i$ are related to the true points X_i by

$$X_i = H\hat{X}_i$$

Where H is an homography with 15 degrees of freedom.

The reconstruction obtained in this way does not preserve angles, distances or parallel lines and it is denoted projective reconstruction

Reconstruction: uncalibrated cameras (2)

Algorithm

- estimate matrix F
- define a pair of cameras compatible with F, e.g.,

$$P = [I \quad 0] \qquad \mathrm{P}' = \left[[\mathrm{e}']_x F \quad e' \right]$$

- use standard stereo reconstruction methods to estimate X

Reconstruction: uncalibrated cameras (3)

Metric reconstruction:

Knowing pairs of corresponding points $(X_i, \hat{X}_i)$

$$Mh = 0$$

$$M = \begin{bmatrix} X^{1T} & 0 & 0 & -x_1^1 X^{1T} \\ 0 & X^{1T} & 0 & -x_2^1 X^{1T} \\ 0 & 0 & X^{1T} & -x_3^1 X^{1T} \\ \vdots & \vdots & \vdots & \vdots \\ X^{nT} & 0 & 0 & -x_1^n X^{nT} \\ 0 & X^{nT} & 0 & -x_2^n X^{nT} \\ 0 & 0 & X^{nT} & -x_3^n X^{nT} \end{bmatrix}$$

$$h = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix}$$

$$H = \begin{bmatrix} h_1^T \\ h_2^T \\ h_3^T \\ h_4^T \end{bmatrix}$$

Referências

R. Hartley, A. Zisserman, Multiple View Geometry, Cambridge University Press, 2000.

D. Forsyth, Ponce, Computer Vision: a Modern approach, Prentice Hall,

P. H. S. Torr, A Structure and Motion Toolkit in Matlab in Interactive Adventures in S and Ml, Microsoft Research, http://research.microsoft.com/~philtorr/, June 2002

CV-Online